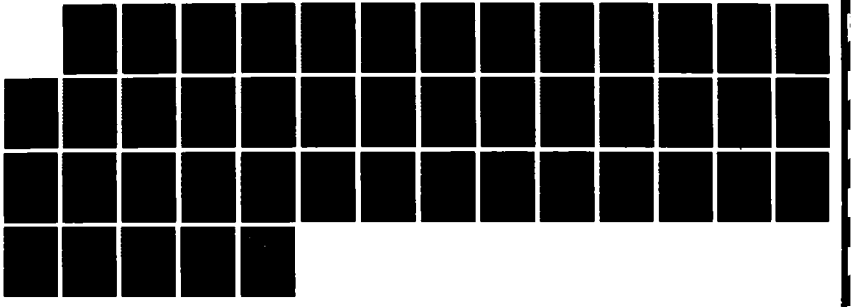
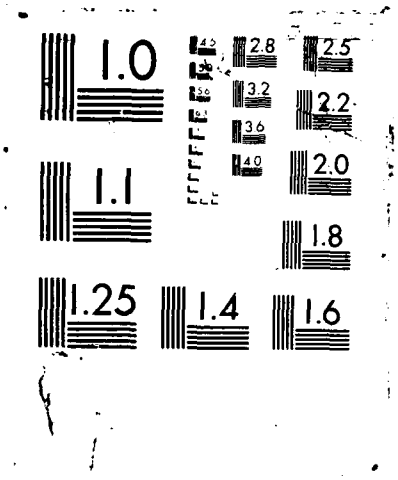
AD-A193 150 DISTRIBUTED ALGORITHMS FOR PROBABILISTIC SOLUTION OF COMPUTATIONAL VISION PROBLEMS(U) SCIENTIFIC SYSTEMS INC CAMBRIDGE MA D E GUSTAFSON ET AL. 01 MAR 88 1/1

UNCLASSIFIED N00014-87-C-0763 F/G 9/5 NL




1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

SSI Project No. 1136

Title: Distributed Algorithms for Probabilistic Solution of Computational Vision Problems

Final Report

Contract No. N00014-87-C-0763

Contractor:

Scientific Systems Inc.
One Alewife Place
Cambridge, MA 02140

Tel: 617-661-6364

Principal Investigator:
Donald E. Gustafson

March 1, 1988

License Rights Legend

Contract No. N00014-87-C-0763
Contractor or Subcontractor: Scientific Systems, Inc.

For a period of two (2) years after the delivery and acceptance of the last deliverable item under this contract, this technical data shall not, without the written permission of the above Contractor, be either (A) used, released or disclosed in whole or in part outside the Government, (B) used by a party other than the Government for manufacture, or (C) used by a party other than the Government. After the expiration of the two (2) year period, the Government may use, duplicate, or disclose the data, in whole or in part and in any manner, for Government purposes only, and may have or permit others to do so for Government purposes only. All rights to use or duplicate the data in whole or in part for commercial purposes are retained by the Contractor, and others to whom this data may be disclosed agree to abide this commercial purposes limitation. The Government assumes no liability for use or disclosure of the data by others for commercial purposes. This legend shall be included on any reproduction of this data, in whole or in part.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

2

REPORT DOCUMENTATION PAGE

ADA193150

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: Scientific Systems, Inc. / 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: Defense Contract Administration Services, SCD-A
6c. ADDRESS (City, State, and ZIP Code): One Alewife Place, Cambridge, MA 02140	7b. ADDRESS (City, State, and ZIP Code): 495 Summer Street, Boston, MA 02210-2184
8a. NAME OF FUNDING / SPONSORING ORGANIZATION: Office of Naval Research / 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N00014-87-C-0763
8c. ADDRESS (City, State, and ZIP Code): 800 N. Quincy Street, Arlington, VA 22217	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO / PROJECT NO / TASK NO / WORK UNIT ACCESSION NO

11. TITLE (Include Security Classification)

Distributed Algorithms for Probabilistic Solutions of Computational Vision Problems (u)

12. PERSONAL AUTHOR(S)

Donald E. Gustafson and Sanjoy K. Mitter

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM 8/87 TO 1/88	88/3/1	

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A new approach is developed for solving the moving target detection and tracking problem using highly cluttered images. The unknown target is assumed to be moving over a cluttered background in the presence of foreground noise. Using a Markov random field model for the target and a probabilistic description of the noise, the posterior distribution of the target is a Gibbs distribution. The maximum aposteriori target image is found by a randomized search process. Both batch and recursive formulations are developed, with the recursive approach yielding superior results. Numerical results indicate that this approach can successfully detect and track small targets in environments where the target is essentially made invisible by noise. The algorithms are almost completely parallelizable: for n pixels a total of n/4 processors may be used, with the result that solutions would require on the order of 2 seconds on current machines for the examples presented. (Keywords: motion; optical flow).

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code) / 22c. OFFICE SYMBOL

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

Contents

Section		Page
1.	INTRODUCTION	1
	1.1 Computational Vision Systems	2
	1.2 Standard Regularization in Early Vision	3
	1.3 Limitations of Standard Regularization Theory	4
	1.4 Stochastic Approach to Regularizing Early Vision	5
	1.5 Phase I Technical Objectives	6
2.	IMAGES AS FUNCTIONS OF MARKOV RANDOM FIELDS	7
3.	ESTIMATING STRUCTURE FROM MOTION	10
	3.1 Optical Flow	11
	3.2 Batch Processing Formulation	12
	3.3 Recursive Problem Formulation	23
	3.4 Parallel Implementations	32
4.	SUMMARY AND CONCLUSIONS	36
	REFERENCES	38

Figures

Figure		Page
3.1	Detection of Moving Object in Cluttered Background and Noise	16
3.1	(Cont'd) Detection of Moving Object in Cluttered Background and Noise	17
3.2	Noise - Free Images of Moving Object ($v = (1,1)$)	18
3.3	True Object in Original Position (k=1) and Cluttered Background	19
3.4	Noisy Images Containing Target Over Cluttered Background with Additive Foreground Noise	20
3.5	Object Detection (at original position) for Velocity Match ($\hat{v} = (1,1)$)	21
3.6	Object Detection for Velocity Mismatch ($\hat{v} = (1,0)$)	21
3.7	True Object at Original Location, and Cluttered Background	29
3.8	Successive Cluttered Images Containing Object over Cluttered Background, with Additive Noise	30
3.9	Object Recovery	31
3.10	Cluttered Images and Background with Increased Noise ($\sigma_B = 1$, $N_0 = 1$, $N_1 = 0.25$)	33
3.11	Object Recovery with Increased Noise	34

1. INTRODUCTION

With the continuing advances in computing hardware and the advent of dedicated parallel processors, it is possible to place more and more computing power in small devices. This increased capability is particularly important for real-time image processing applications which require extremely high throughput. Examples of such application areas are robotic vision, medical diagnosis and military target acquisition and tracking, such as SDI. Current algorithms tend to be ad hoc in nature, typically consisting of a cascade of processors which in some cases work at cross purposes. For example, a 2D filter is often used to reduce noise with high spatial frequency content. However, this operation tends to blur edges which must be detected later to determine object boundaries. A more suitable approach would be to combine edge detection and noise reduction in a single step, if possible. Such an approach would have the potential for increased performance, particularly at low signal/noise ratios, where false alarm and miss rates tend to increase rapidly in current systems.

Future image processing systems should be able to solve a variety of image understanding problems, such as:

- image segmentation
- surface reconstruction
- stereo matching
- determining structure from motion.

The field of computational vision is dedicated to solving these types of problems and has been developing rapidly over the past fifteen years.

Within the last four years, some exciting developments have occurred which show great promise in providing coordinated solutions to these problems utilizing distributed algorithms. These developments are based on utilizing a probabilistic framework. The two-dimensional image is modeled as a random field which has to be estimated in real time from a set of noisy ambiguous measurements from multiple sensors. A Bayesian viewpoint is adopted, in which the prior knowledge is expressed as a probability distribution. Using a probabilistic description of the observation noise, the posterior distribution of the random field can be computed. These models are based on the use of Markov random fields and the Gibbs distribution. Significantly, these assumptions lead to distributed algorithms which may be implemented on parallel processors. There are several other important advantages in using this approach (Morroquin, Mitter and Poggio, 1986). It is possible to model both piecewise continuous surfaces and the boundaries between smooth patches (targets, clouds, objects, e.g.). It provides a general framework for solving all of the problems mentioned above. The parameters that appear in the reconstruction algorithms have a precise statistical interpretation which may be validated on physical grounds.

1.1 Computational Vision Systems

The standard definition of computational vision is that it is inverse optics. The direct problem-the problem of classical optics-or computer graphics-is to determine the images of three-dimensional objects. Computational vision is confronted with inverse problems of recovering surfaces from images. Much information is lost during the imaging process

that projects a three-dimensional world into two-dimensional arrays (images). As a consequence, vision must rely on natural constraints, that is, general assumptions about the physical world to derive an unambiguous output. This is typical of many inverse problems in mathematics and physics.

The first part of vision - from images to surfaces - has been called early vision. The common characteristics of most early vision problems, in a sense their deep structure, can be formalized: early vision problems are ill-posed in the sense defined by Hadamard (Poggio and Torre (1984)). A problem is well-posed when its solution (a) exists, (b) is unique and (c) depends continuously on the initial data. Ill-posed problems fail to satisfy one or more of these criteria.

1.2 Standard Regularization in Early Vision

The main idea for "solving" ill-posed problems is to restrict the class of admissible solutions suitable a priori knowledge. In standard regularization methods, due mainly to Tikhonov, the regularization of the ill-posed problem of finding z from the data y: $Az = y$ requires the choice of norms $||\cdot||$ and of a stabilizing functional $||Pz||$. In standard regularization theory, A is a linear operator, the norms are quadratic and P is linear. A method that can be applied is:

Find z that minimizes $||Az - y||^2 + \lambda||Pz||^2$, where λ is a so-called regularization parameter.

In this method, λ controls the compromise between the degree of regularization of a solution and its closeness to the data. P embeds the phy-

sical constraints of the problem. It can be shown for quadratic variational principles that under mild conditions the solution space is convex and a unique solution exists.

Poggio et al (1984, 1985) show that several problems in early vision can be "solved" by standard regularization techniques. Surface reconstruction, optical flow at each point in the image, optical flow along contours, color, stereo can be computed by using standard regularization techniques. Variational principles that are not exactly quadratic but have the same form as that above can be used for other problems in early vision. The main results of Tikhonov can, in fact, be extended to some cases in which the operators A and P are nonlinear, provided they satisfy certain conditions. (Morozov, 1984.)

1.3 Limitations of Standard Regularization Theory

Standard regularization theory with linear A and P is equivalent to restricting the space of solution to generalized splines, whose order depends on the order of the stabilizer P. This means that in some cases the solution is too smooth, and cannot be faithful in locations where discontinuities are present. In optical flow, surface reconstruction and stereo, discontinuities are in fact not only present, but also the most critical locations for subsequent visual information processing. Standard regularization cannot deal well with another critical problem of vision, the problem of fusing information from different early vision modules. Since the regularizing principles of the standard theory are quadratic, they lead to linear Euler-Lagrange equations. The output of different

modules can therefore be combined only in a linear way. Terzopoulous (1984; see also Poggio et al., 1985) has shown how standard regularization techniques can be used in the presence of discontinuities in the case of surface interpolation. After standard regularization, locations where the solution f originates a large error in the regularization term, are identified (this needs setting a threshold for the error in smoothness). A second regularization step is then performed using the location of discontinuities as boundary conditions.

In any case, one would like a more comprehensive and coherent theory capable of dealing directly with the problem of discontinuities and the problem of fusing information. So the challenge for a regularization theory of early vision is to extend it beyond standard regularization methods and their most obvious non-linear versions.

1.4 Stochastic Approach to Regularizing Early Vision

In this research, we have developed a rigorous approach to overcome part of the ill-posedness of vision problems, based on Bayes estimation and Markov Random Field models, that effectively deals with the problems faced by the standard regularization approach. In this approach, the a priori knowledge is represented in terms of an appropriate probability distribution, whereas in standard regularization a priori knowledge leads to restrictions on the solution space. This distribution, together with a probabilistic description of the noise that corrupts the observations, allows one to use Bayes theory to compute the posterior distribution $P_{f|g}$,

which represents the likelihood of a solution f given the observations g. In this way, we can solve the reconstruction problem by finding the estimate $\hat{f}$ which either maximizes this a posteriori probability distribution (the so called Maximum a Posteriori or MAP estimate), or minimizes the expected value (with respect to $P_{f|g}$) of an appropriate error function. The class of solutions that can be obtained in this way is much larger than in standard regularization.

The price to be paid for this increased flexibility is computational complexity. New parallel architectures and possibly hybrid computers of the digital-analog type promise however to deal effectively with the computational requirements of the methods proposed here.

We wish to emphasize here that our main thrust here is in development of distributed algorithms suitable for parallel architectures, and on comprehensive testing on image data.

1.5 Phase I Technical Objectives

Our research objectives were to:

- develop new distributed algorithms for recovering structure from motion and for discrimination of known or unknown objects from highly cluttered background.
- assess feasibility of real-time operation using state-of-the-art parallel processors
- evaluate performance using highly-cluttered image data containing moving targets.

2. IMAGES AS FUNCTIONS OF MARKOV RANDOM FIELDS

The key to success in the use of the proposed approach, is the ability to find a class of stochastic models (that is, random fields) that have the following characteristics:

(i) The probabilistic dependencies between the elements of the field should be local. This condition is necessary if the field is to be used to model surfaces that are only piecewise smooth; besides, if it is satisfied, the reconstruction algorithms are likely to be distributed, and thus, efficiently implementable in parallel hardware.

(ii) The class should be rich enough, so that a wide variety of qualitatively different behaviors can be modeled.

(iii) The relation between the parameters of the models and the characteristics of the corresponding sample fields should be relatively transparent, so that the models are easy to specify.

(iv) It should be possible to represent the prior probability distribution P_f explicitly, so that Bayes theory can be applied.

(v) It should be possible to specify efficient Monte Carlo procedures, both for generating sample fields from the distribution, so that the capability of the model to represent our prior knowledge can be verified, and to compute the optimal estimators.

A class of random fields that satisfies these requirements is the class of Markov Random Fields (MRF's) on finite lattices. A MRF has the property that the probability distribution of the configurations of the field can always be expressed in the form of a Gibbs distribution:

$$P_f(f) = \frac{1}{Z} \exp[- \frac{1}{T_0} E(f)] \tag{2.1}$$

where Z is a normalizing constant, T_0 is a parameter (known as the "natural temperature" of the field) and the "Energy function" E(f) is of the form:

$$E(f) = \sum_C V_C(f)$$

where C ranges over the "cliques" associated with the neighborhood system of the field, and the potentials $V_C(f)$ are functions supported on them (a clique is either a single site, or a set of sites such that any two sites belonging to it are neighbors of each other).

We will assume that the available observations g are obtained from a typical realization f of the field by a degrading operation (such as sampling) followed by corruption with i.i.d. noise (the form of whose distribution is known), so that the conditional distribution can be written as:

$$P_{g|f}(g;f) = \exp[-\alpha \sum_{i \in S} \phi_i(f,g_i)]$$

where $\{\phi_i\}$ are some known functions, and α is a parameter. The posterior distribution is obtained from Bayes rule:

$$P_{f|g}(f;g) = \frac{1}{Z_p} \exp[-E_p(f;g)] \tag{2.2}$$

$$E_p(f;g) = \frac{1}{T_0} E(f) + \alpha \sum_{i \in S} \phi(f,g_i) \tag{2.3}$$

It is important to note that the Markov structure is retained under conditioning and that the posterior distribution is also a Gibbs distribution.

Cost Functionals

The Bayesian approach to the solution of reconstruction problems has been adopted by several researchers. In most cases, the criterion for selecting the optimal estimate has been the maximization of the posterior probability (the Maximum aposteriori or MAP estimate). It has been used, for example, by Geman and Geman (1984) for the restoration of piecewise constant images; by Grenander (1984) for pattern recognition, and by Elliot et. al. (1983) and Hansen and Elliot (1982) for the segmentation of textured images (a similar criterion - the maximization of a suitably defined likelihood function - has been used by Cohen and Cooper (1984) for the same purposes).

In some other cases, a performance criterion, such as the minimization of the mean squared error has been implicitly used for the estimation of particular classes of fields. For example, for continuous-valued fields with exponential autocorrelation functions, corrupted by additive white Gaussian noise, Nahi and Assefi (1972) and Habibi (1972) have used causal linear models and optimal (Kalman) linear filters for solving the reconstruction problem.

Although other criteria are possible (cf Morroquin, 1985), we have chosen the MAP criterion here for designing optimal estimators since this criterion gives generally equivalent results, with the exception of very high noise situations. The performance of estimators using other criteria would be an appropriate topic for Phase II research.

3. ESTIMATING STRUCTURE FROM MOTION

Since the primary goal of image processing will be target detection, discrimination and monitoring, we can exploit the fact that the target image will be moving relative to most of the rest of the image.

A common approach to the problem is the use of flow fields (for example, Ullman, 1981; Bruss and Horn, 1981; Williams, 1981, Hildreth, 1984; Adiv, 1984). These approaches generally assume deterministic models or utilize standard regularization techniques. We have already alluded to several problems encountered when using these approaches and the potential advantages of a statistically-based approach based on local interaction models.

Reed, et al (1983) have developed a three-dimensional matched filtering approach to moving target detection. However, it is limited to point targets. Legters and Young (1982) developed an operator-based approach using foreground and background models and solved a least-squares minimization problem. No statistical object model was used. Miller, et al (1985) have studied the general moving target detection problem, restricted to point targets, and concluded that the uniformly most powerful detector, invariant with respect to image intensity variations, consists of specific spatial-temporal differencing schemes. In the sequal, we derive similar conclusions for a much more general object model.

We suggest that the MRF methodology may be employed to recover object motion from successive images. The key to the approach lies in appropriately defining the potential function.

3.1 Optical Flow

Let $v_x(x,y)$, $v_y(x,y)$ be the components of the velocity vector at the point (x,y) on the image. Then v_x and v_y can be estimated by using a flow equation

$$M(x,y,t) = v_x \frac{\partial f(x,y)}{\partial x} + v_y \frac{\partial f(x,y)}{\partial y} + \frac{\partial f(x,y)}{\partial t} = 0$$

where $\partial f(x,y)/\partial x$ and $\partial f(x,y)/\partial y$ are spatial gradients of the image intensity at (x,y) and $\partial f(x,y)/\partial t$ is found by time differencing successive images. All terms are readily estimated by using numerical differencing methods. The solution to this equation is a locus of points along a straight line. By evaluating solutions around a neighborhood of (x,y), one can determine estimates for $v_x(x,y)$ and $v_y(x,y)$ by the intersection of the individual solutions over pixels within a neighborhood.

Flow field methods may not be optimal for the SDI problem, however. The critical aspect of the tracking problem we are addressing is the ability to handle cluttered background and foreground noise at very low signal/noise ratios. Methods to date are generally based on moving target indicator (MTI) technology and use flow field analysis. These methods assume smoothness properties which do not hold for objects covering only a few pixels, when only a few grey levels are used, or for the case when object velocities are on the order of one pixel/sample, which is likely to occur in an SDI enviorment. Differencing operations are required which can lead to large errors in high noise conditions. In these cases, detection and tracking accuracy will be degraded by the effects of incorrect assumptions in the problem formulation.

The approach we are studying avoids these problems by analyzing the image on a pixel-by-pixel basis, with no impicit smoothness assumptions. Smoothness may be employed explictly as required, for example to develop a local Markov random field model.

3.2 Batch Processing Formulation

We are interested in recovering the intensity field of an object moving in a highly cluttered environment. The object is modeled using a Gibbs distribution of the form

$$P_f = \frac{1}{z} \exp[-V_c(f)/T] \tag{3.1}$$

where z is a normalizing constant, f is the object intensity field, T is the "temperature", and $V_c(f)$ is an appropriately chosen potential function. We have experimented with several different potential functions, including the Ising model, with the result that the following function

$$V_c(f) = \sum_{N_i} | f_i - f_j | , \tag{3.2}$$

where N_i is the neighborhood of the $i^{\underline{th}}$ pixel and f_i is the intensity of the $i^{\underline{th}}$ pixel, works reasonably well. In our research we have used a 4 - connected neighborhood consisting of the 4 adjacent pixels. The probability density P_f describes the apriori information.

In order to process cluttered images we need a stochastic image model which accounts for object motion, background clutter, and image noise. We have used the following model for the observed image g at time step k:

$$g_i(k) = b_i[1-s(f_{i_k})] + f_{i_k} + n_i(k) \tag{3.3}$$

where b_i is the (fixed) background, i_k indexes the pixels according to the velocity of the object, and s(.) is a function which is one if the object covers pixel i_k at time step k and is zero otherwise. For a two-dimensional field indexed by p,q:

$$p_k = p - k\ v_x$$

$$q_k = q - k\ v_y$$

where v_x, v_y are orthogonal velocity components. Note that we make no smoothness assumptions such as is done in visual flow field analysis. This allows us to analyze images with only a relatively few number of pixels. The noise $n_i(k)$ consists of two processes

$$n_i(k) = n_{oi} + n_{ki} \tag{3.4}$$

where n_{oi} is a time-invariant noise present at all time steps and n_{ki} is a time-varying noise. Both processes are assumed to be exponentially distributed with known mean and variances and are mutually independent. By assuming zero-mean Gaussian distributions, the problem becomes one of minimizing the following energy function with respect to b and f:

$$E(f,g) = V_c(f) / T$$

$$+ \frac{1}{N} \sum_{i,k} \left(g_i(k) - b_i\ [1-s(f_{i_k})] - f_{i_k} \right)^2 \tag{3.5}$$

where N is the sum of the variances of the two noise processes.

The minimization was carried out as follows:

(i) An estimate of b_i was calculated as : $\hat{b}_i = \hat{b}$ with $\hat{b}$ the mean image intensity over the entire image field.

(ii) Select a pixel i at random.

(iii) Select a new estimate $\hat{f}_i \neq f_i$ from the set of allowable intensities using a uniformly distributed random variable.

(iv) Compute the resulting change in energy ΔE.

(v) If $\Delta E < 0$, accept the change, i.e., set $f_i = \hat{f}_i$.

(iv) If $\Delta E > 0$, accept the change if $\exp(-\Delta E) > r$, where r is a uniformly distributed random variable between 0 and 1; otherwise, leave f_i unchanged.

(iiv) Repeat steps (ii) - (vi).

Steps (ii) - (vi) comprise the Metropolis algorithm (Metropolis, et al 1953), which generates a regular, reversible Markov chain. The resulting image configurations are distributed according to the Gibbs distribution

$$\frac{1}{z} \exp [- E(f, g)]$$

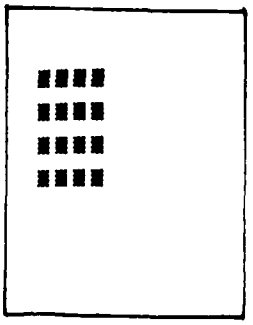
Minimization with respect to velocity was performed by an exhaustive search over all possible velocities, using a course quantization. Note that this minimization can be done using completely independent processors.

We have conducted experiments on several different cases, with promising results. An example is shown in Figure (3.1), where a square 4x4 object is moving at the rate of one pixel/sample to the right over a cluttered 10x14 background and is viewed in noise. The object has intensity of 2 over a background whose mean intensity is 1 with an rms value of 0.5. The additive noise has an rms value of 0.5 and T = 1. Four intensity levels (0,1,2,3) are plotted. Figure (3.1) shows the actual object, actual background (panel b) and 4 successive cluttered images of the moving object (panels c-f). The object has essentially disappeared in the noise. Recovery is shown in panels g-k: the object is fully recovered using only 16 iterations/pixel (we are assumung parallel implementation). The error shown is $\sum_i (f_i - f_i)^2$ while the energy shown is E for a temperature of 1. The initial energy (using as b the mean pixel intensity) was 601.017. The minimum energy solution clearly recovers the original object. Note that this case corresponds to recovering the object in the time required for it to move by one length.

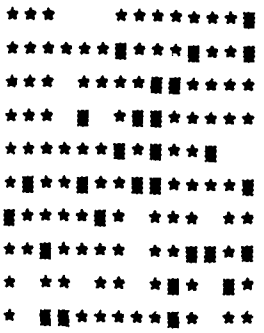
The previous tests used the Metropolis algorithm to detect moving objects. We have found that the algorithms may be simplified considerably without any significant loss of accuracy by simplifying the iterations. The simplification is to only accept changes which lead to a reduction in the energy fuction.

An example of our results is given in Figures 3.2 - 3.6. In Figure 3.2, a series of six noise-free images are shown in which a 3x3 object of

Figure 3.1 Detection of Moving Object in Cluttered Background and Noise

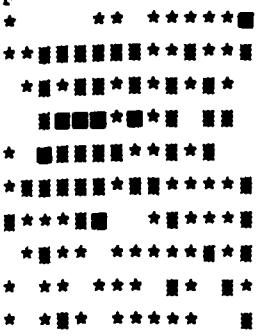


(a) noise-free image (object) at k=1, velocity is one pixel/sample to the right



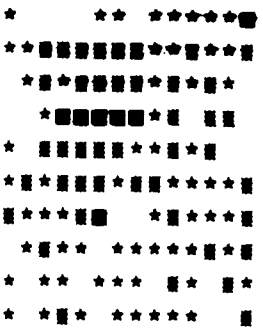
(b) actual background

phi:k= 1



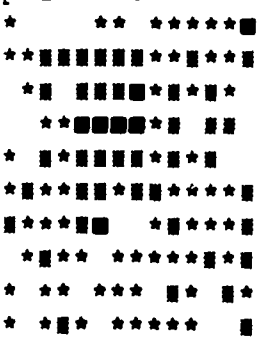
(c) noisy image (k=1)

phi:k= 2



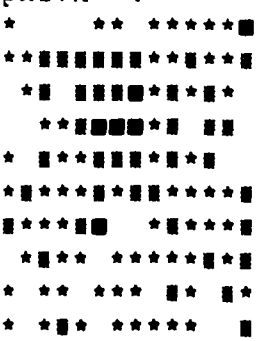
(d) noisy image (k=2)

phi:k= 3



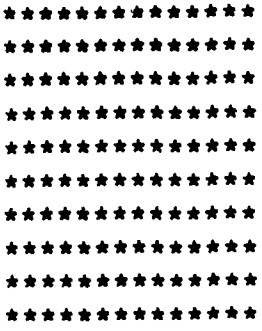
(e) noisy image (k=3)

phi:k= 4

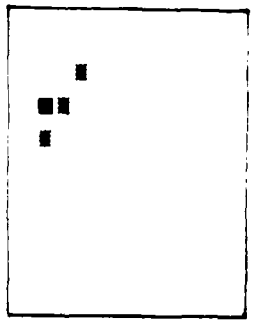


(f) noisy image (k=4)

Figure 3.1 (cont'd) Detection of Moving Object in Cluttered Background and Noise

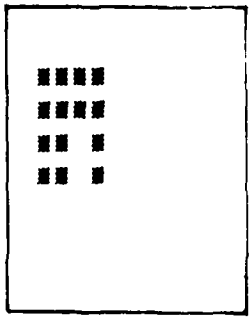


(g) estimated background



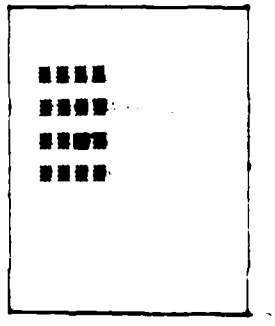
energy 607.3359
error 49

(h) object estimate
(1 iteration/pixel)



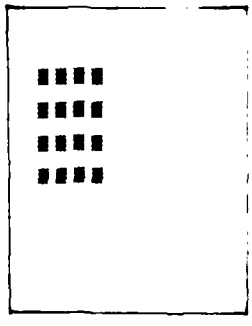
energy 563.8228
error 8

(i) object estimate
(6 iterations/pixel)



energy 552.8715
error 1

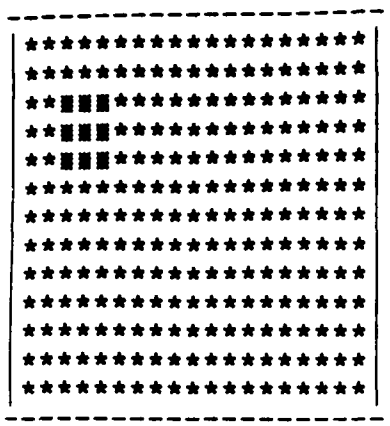
(j) object estimate
(11 iterations/pixel)



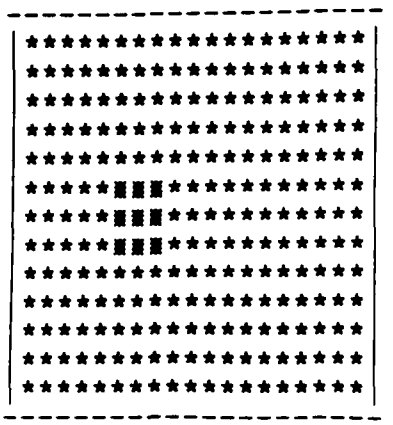
energy 538.2726
error 0

(k) object estimate
(16 iterations/pixel)

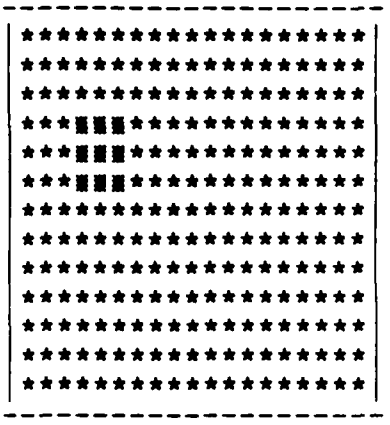
phi:k= 1



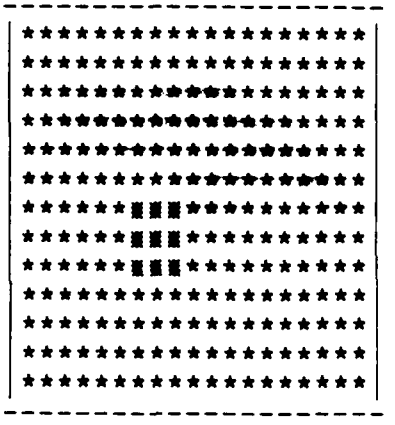
phi:k= 4



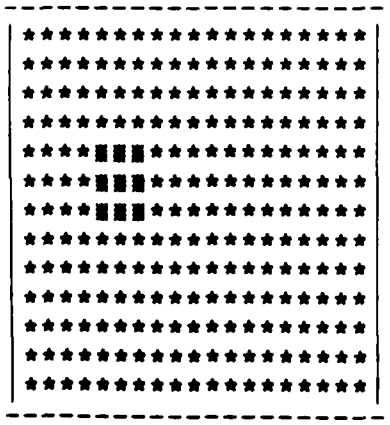
phi:k= 2



phi:k= 5



phi:k= 3



phi:k= 6

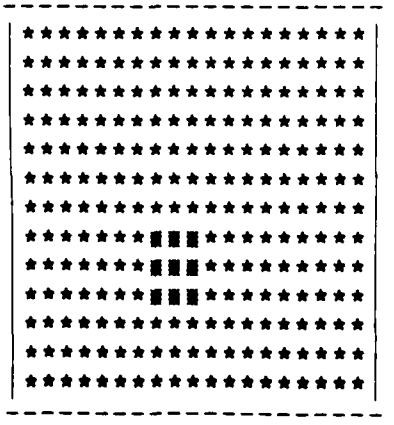
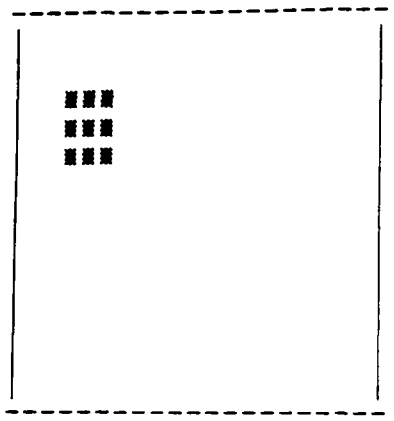


Figure 3.2 Noise - Free Images of Moving Object (v = (1,1)).

(a) psitru



(b) background(btru)

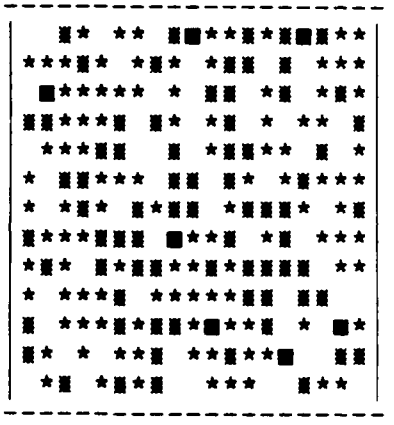
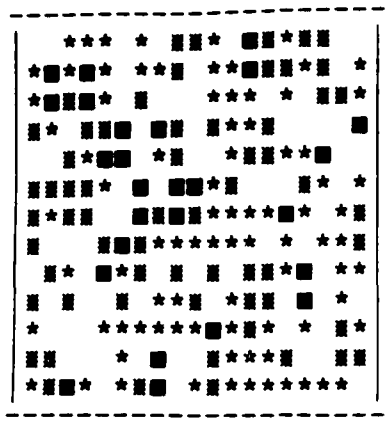
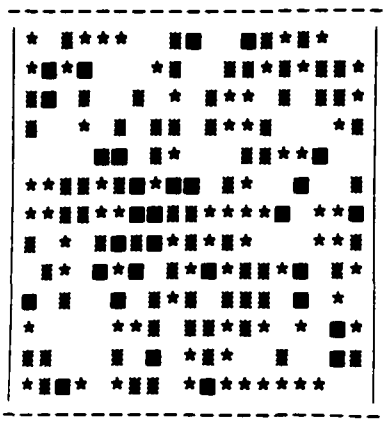


Figure 3.3 True Object in Original Position (k=1) and Cluttered Background.

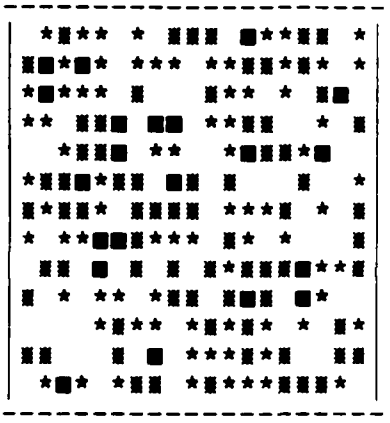
phi:k= 1



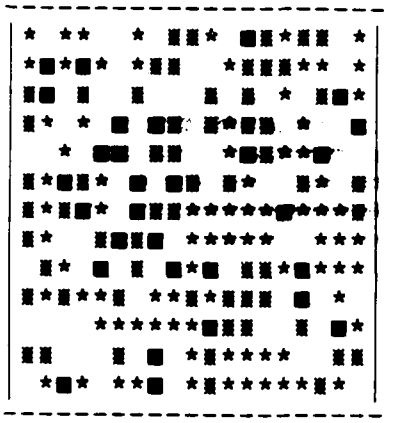
phi:k= 4



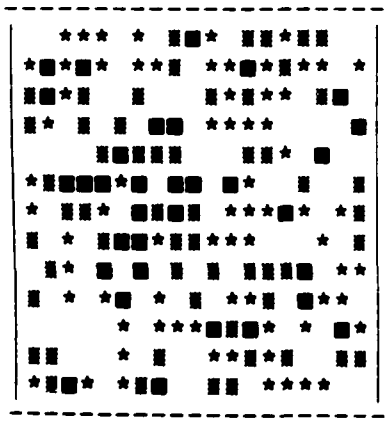
phi:k= 2



phi:k= 5



phi:k= 3



phi:k= 6

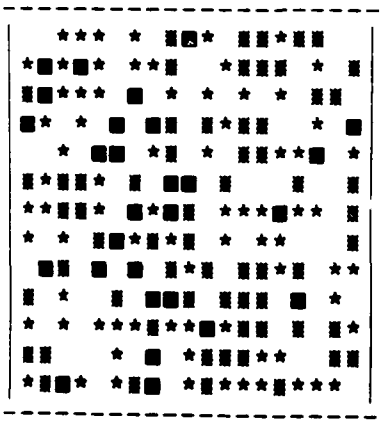


Figure 3.4 Noisy Images Containing Target Over Cluttered Background with Additive Foreground Noise.

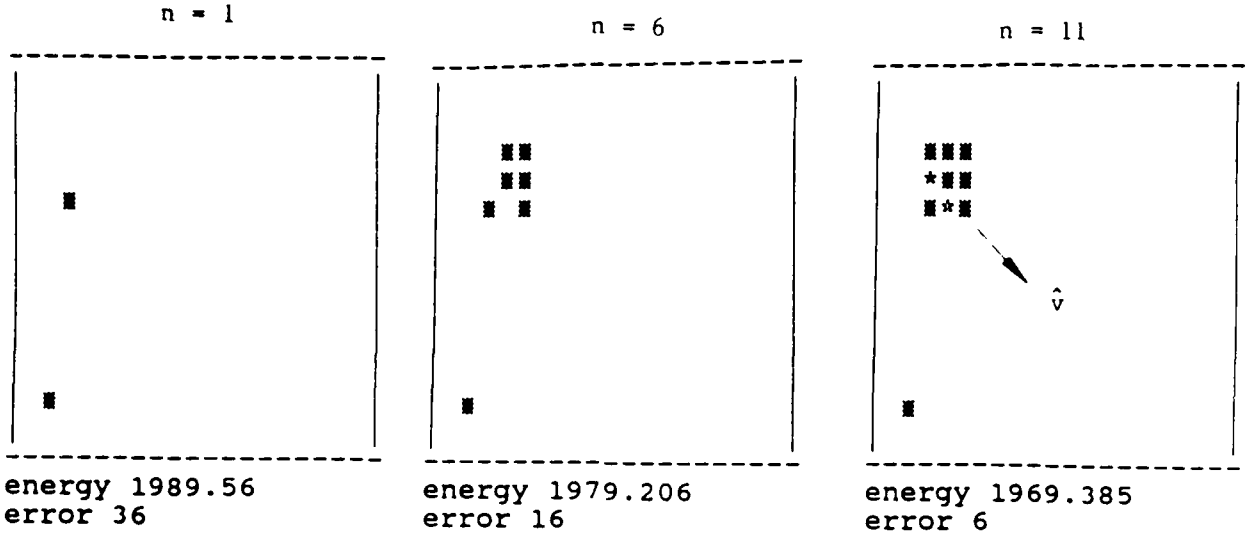


Figure 3.5 Object Detection (at original position) for Velocity Match ($\hat{v}$ = (1,1)).

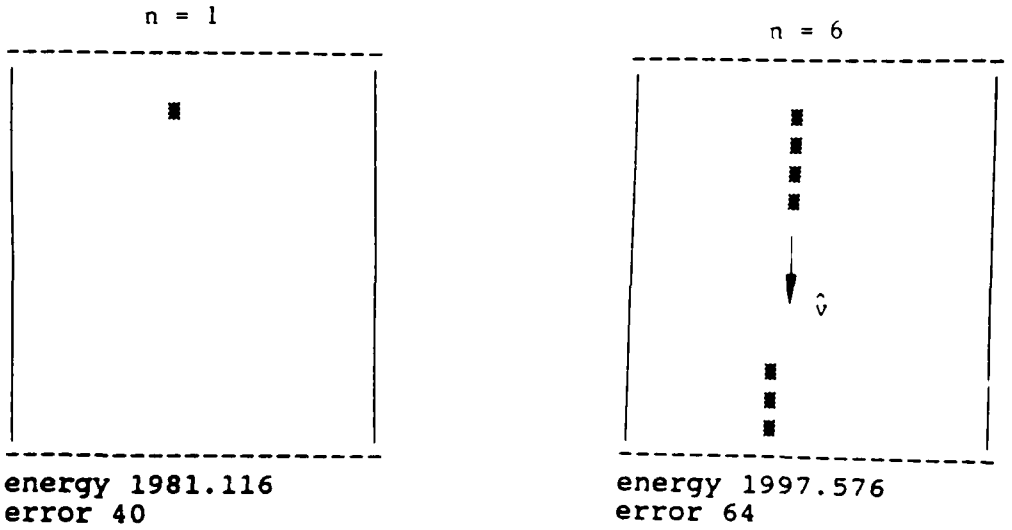


Figure 3.6 Object Detection for Velocity Mismatch ($\hat{v}$ = (1,0)).

intensity 2 moves over a 13x19 background of intensity 1 with velocity v = (1,1) pixel/sample. The object in its original position over a background of intensity 0 is shown in Figure 3.3. In Figure 3.3, a cluttered background is shown which was generated using Gaussian random noise with rms value of 0.8. The intensities are quantized in the figure to four values (0,1,2,3). Noisy images were then generated by adding time-invariant noise (rms = 0.8) and time-varying noise (rms = 0.4) to the image comprised of the object over the cluttered background. The resulting images are shown in Figure 3.4; the object has essentially disappeared in the noise. These six images were used in the detection algorithm to search for the presence of a moving object. The results are shown in Figure 3.5 and 3.6. In Figure 3.5, the estimated velocity matches the actual velocity and the object is recovered, with only small error and is in the correct position. Three cases are shown for n=1, 6, 11 where n is the number of global iterations (a global iteration is comprised of one iteration per pixel over the entire image).

In Figure 3.6, the results for the case of a velocity mismatch ($\hat{v}$ = (1,0)) are shown: two long slender objects are detected which are aligned along the estimated velocity direction and are in the wrong position. However, the energy is higher than for the case of a velocity match, indicating that there is a velocity error. Thus, the minimum energy solution yields the correct velocity, as desired. The error values shown in the figures are the true (unobservable) fit error measured in terms of the pixel-by-pixel difference between the object estimate and the true object.

3.3 Recursive Problem Formulation

During the course of our research we developed a new problem formulation which yields improved results over the previous batch processing approach. It is based on a recursive Bayes formulation to compute the object conditional density, based on the observed images to date.

In our work we have used the following model:

$$\Phi_k(i) = B(i)\,[1-S(\Psi(i_k)] + \Psi(i_k) + n_k(i) + n_0(i) \qquad (3.6)$$

$$k = 1,2, \ldots .$$
$$i = 1,2, \ldots .$$

where

$\Phi_k(i)$ = observation intensity at pixel i and time k

$\Psi(i_k)$ = object intensity at pixel i_k and time k. The index i_k depends upon the motion of the object. Throughout we will set $i_1 = i$.

$B(i)$ = background intensity at pixel i

$n_0(i)$ = time-invariant foreground noise

$n_k(i)$ = time-varying foreground noise

$$S(\Psi(i_k)) = \begin{cases} 0 \ ; \ \Psi(i_k) = 0 \\ 1 \ ; \ \text{otherwise} \end{cases} \qquad (3.7)$$

Note that B(i) is time-invariant. Also, without loss of generality we take $n_1(i) = 0$. We will assume that $n_k(i)$ and $n_0(i)$ are mutually independent gaussian zero mean processes, uncorrelated over both space and time, with

$$E\,[n_k(i)^2] = N_1 \text{ for all } k$$

$$E\,[n_0(i)^2] = N_0$$

Let $\Psi = \{\Psi(i)\}$

$\Phi_k = \{\Phi_k(i)\}$

$B_k = \{B(i)\}$

Then we are interested in the probability

$$P(\Psi,\ B \mid \Phi_1,\ \Phi_2,\ \ldots,\ \Phi_n)$$

By Bayes' Rule we get the recursion

$$P(\Psi,\ B \mid \Phi_1,\ \ldots,\ \Phi_n) = \frac{P(\Phi_n \mid \Psi,\ B,\ \Phi_1,\ \ldots,\ \Phi_{n-1})\ P(\Psi,\ B \mid \Phi_1\ \ldots\ \Phi_{n-1})}{P(\Phi_n \mid \Phi_1\ \ldots\ \Phi_{n-1})}$$

We start with the initial conditions, using as apriori probability

$$P_0(\Psi,\ B) = \frac{1}{z_0}\ \exp\left[\ -\ \frac{1}{\beta_1}\ U_1(\Psi)\ -\ \frac{1}{\beta_2}\ U_2(B)\right]$$

where $U_1(.)$ and $U_2(.)$ are appropriate energy functions. For example, we can utilize Ising potentials. These potentials are completely arbitrary and yield a Gibbs distribution.

Let

$$P_0(\Psi,\ B) = \frac{1}{z_0}\ \exp\ [-\ U_0(\Psi,\ B)]$$

Then

$$P(\Psi, B \mid \Phi_1) = \frac{P(\Phi_1 \mid \Psi, B)\, P_0(\Psi, B)}{P(\Phi_1)}$$

The denominator is independent of Ψ and B. Thus, if we want a point estimate (Maximum Aposteriori e.g.) we need only evaluate the numerator. We will assume we are looking for a MAP estimator here. Neglecting terms independent of Ψ and B we get a simplified form.

Now

$$\Phi_1(i) = B(i)\,[1-S(\Psi(i)] + \Psi(i) + n_0(i)$$

Recall that there is no time-varying noise component here.

We now introduce our motion model. We assume that the object we are interested in tracking is associated with Ψ in the following way:

(a) if $\Psi(i_k) > 0$, then the object is within pixel i_k at time k

(b) if $\Psi(i_k) = 0$, then the object is not within pixel i_k at time k

We further require that $\Psi(i_k) \geq 0$ for all i,k.

We assume here constant rectilinear motion, which is consistent with the space-based target tracking problem over a short time period. Over a longer time period, the tracking algorithm can be used on successive data windows to achieve accurate tracking for non-rectilinear motion. Thus we can write, in two dimensions

$$\Psi(i_k, j_k) = \Psi(i-k\ v_x, j-k\ v_y)$$

where v_x, v_y are the velocity components. This is the discrete form of the optical flow equation. Returning to our development we find that the con-

ditional density for the first image is

$$P(\Phi_1 \mid \Psi, B) = \prod_i P_{n_0} \{\Phi_1(i) - B(i) [1-S(\Psi(i)] - \Psi(i)\}$$

where P_{n_0} is the probability density of n_0. Combining these results, we find that the aposteriori density given the first image may be written as a Gibbs distribution

$$P(\Psi, B \mid \Phi_1) = \frac{1}{z_1} \exp \{- E_1(\Psi_0, B, \Phi,)\}$$

where

$$E_1(\Psi_0, B, \Phi_1) = \frac{1}{2N_0} \sum_i \{\Phi_1(i) - B(i) [1-S(\Psi_0(i_1))] - \Psi_0(i_1)\}^2$$

$$+ U_0(\Psi, B)$$

We now go to the next time step. The second image is modeled as

$$\Phi_2(i) = B(i) [1-S(\Psi_0(i_2)] + \Psi_0(i_2) + n_0(i) + n_2(i)$$

By subtracting $\Phi_1(i)$ we get

$$\Phi_2(i) - \Phi_1(i) = B(i) [-S(\Psi_0(i_2)) + S(\Psi_0(i_1))]$$

$$+ \Psi_0(i_2) - \Psi_0(i_1) + n_2(i)$$

Now

$$P(\Psi, B \mid \Phi_1, \Phi_2) = \frac{P(\Phi_2 \mid \Psi, B, \Phi_1) P(\Psi, B \mid \Phi_1)}{P(\Phi_2 \mid \Phi_1)}$$

and

$$P(\Phi_2 \mid \Psi, B, \Phi_1) =$$

$$\prod_i P_{n_2} \{\Phi_2(i) - \Phi_1(i) + B(i) [S(\Psi(i_2)) - S(\Psi(i))] + \Psi(i) - \Psi(i_2)\}$$

where $P_{n_2}(.)$ is the probability density of n_2.

Thus

$$P(\Psi, B \mid \Phi_2) = \frac{1}{z_2} \exp \{- E(\Psi_0, B, \Phi_1, \Phi_2)\}$$

where

$$E(\Psi_0, B, \Phi_1, \Phi_2) = U_0 (\Psi_0, B)$$

$$+ \frac{1}{2N_0} \sum_i \{\Phi_1(i) - B_i) [1-S(\Psi(i)] - \Psi(i)\}^2$$

$$+ \frac{1}{2N_1} \sum_i \{\Phi_2(i) - \Phi_1(i) + B(i) [S(\Psi(i_2)) - S(\Psi(i))] + \Psi(i) - \Psi(i_2)\}^2$$

We can see immediately by induction that

$$P(\Psi, B \mid \Phi_1, \ldots, \Phi_n) = \frac{1}{z_n} \exp \{- E(\Psi_0, B, \Phi, \ldots \Phi_n)\}$$

where

$$E(\Psi_0, B, \Phi_1, \ldots, \Phi_n) = U_0(\Psi_0, B)$$

$$+ \frac{1}{2N_0} \sum_i \{\Phi_1(i) - B(i) [1-S(\Psi(i))] - \Psi(i)\}^2$$

$$+ \frac{1}{2N_1} \sum_k \sum_i \{\Phi_k(i) - \Phi_1(i) + B(i) [S(\Psi(i_k)) - S(\Psi(i))] + \Psi(i) - \Psi(i_k)\}^2 \quad (3.8)$$

This expression is to be minimized with respect to Ψ_0, B, v_x and v_y. Note that E is nonlinear in Ψ.

Minimization can be done for example using gradient approximations or stochastic approximations or simulated annealing. In our work to date we have used the randomized descent method described in Section 3.3, which is highly parallelizable.

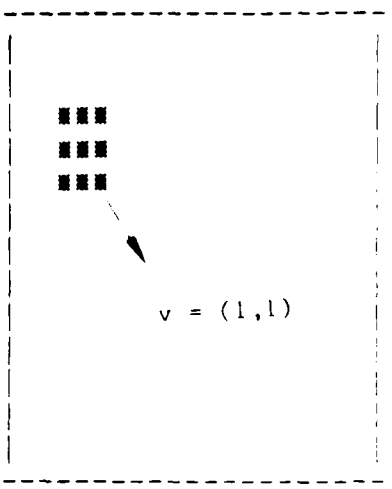
If we have no prior information on B, then $U_0(\Psi_0, B) = U_0(\Psi_0)$. However, if we assume, say, that object is generally brighter that background we can construct an appropriate potential function. Here, we have assumed no prior information on B.

Experimental Results

We present results on the same case as presented in Figures 3.2 - 3.6, in which there was some error in recovering the object. Figure 3.7 shows the object at its original position (k = 1) which is moving with velocity v = (1,1) pixel/sample. Also shown is the background which had mean one and rms 0.8 (gaussian). Figure 3.8 shows successive cluttered images in which the object is moving over the cluttered background in additive noise ($N_0 = 0.64$, $N_1 = 0.16$). Figure 3.9 shows the object recovery for three different values of estimated velocity after n global iteratios. When the velocity matches (panel (a)), the global minimum energy is achieved and the object is recovered exactly. This is an improvement in performance over the results presented in Figure 3.5. Note that no more than 11 global iterations are required for recovery. The indicated error is the true sum-squared pixel error.

Since recovery was perfect for this case, we increased the noise to $\sigma_B = 1.0$, $N_0 = 1$, $N_1 = 0.25$, a decrease of 36% in signal/noise ratio for

psitru



background(btru)

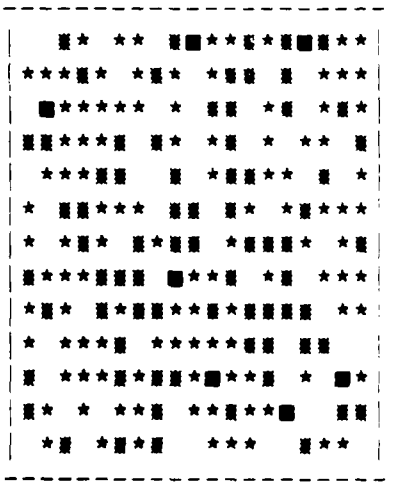


Figure 3.7 True object at Original Location, and Cluttered Background.

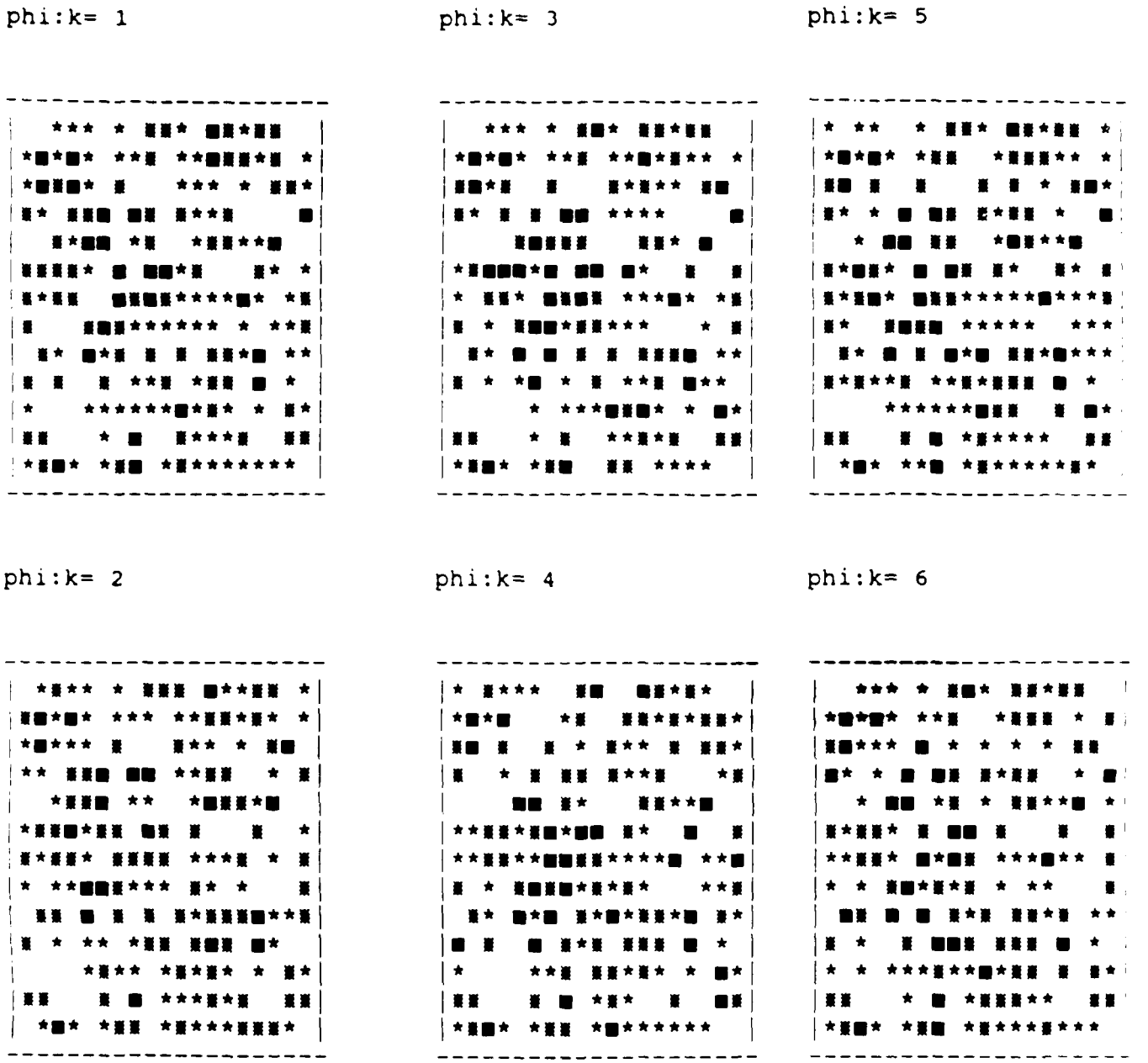
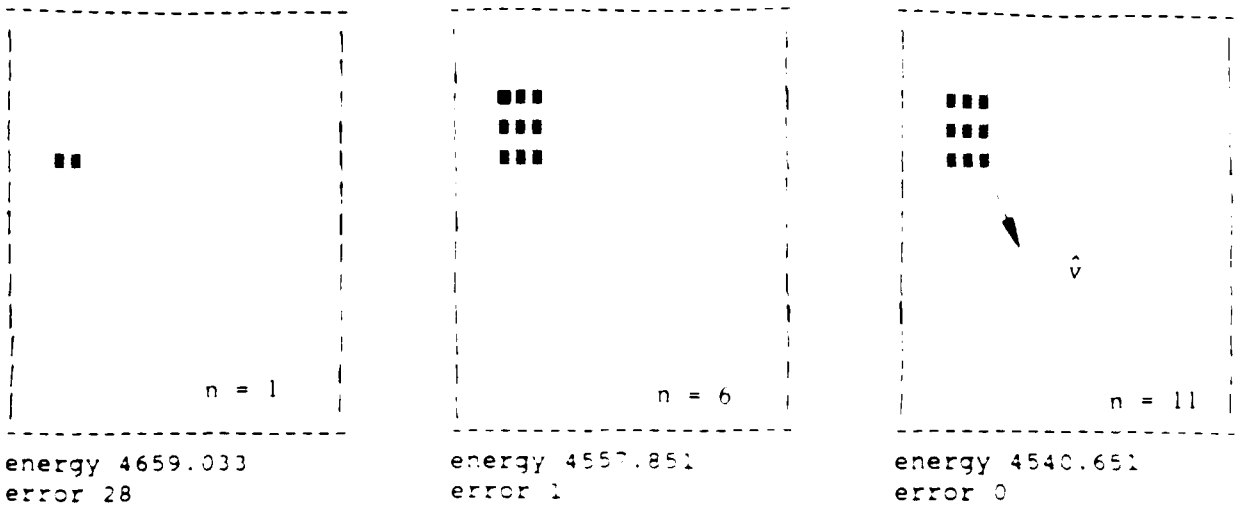
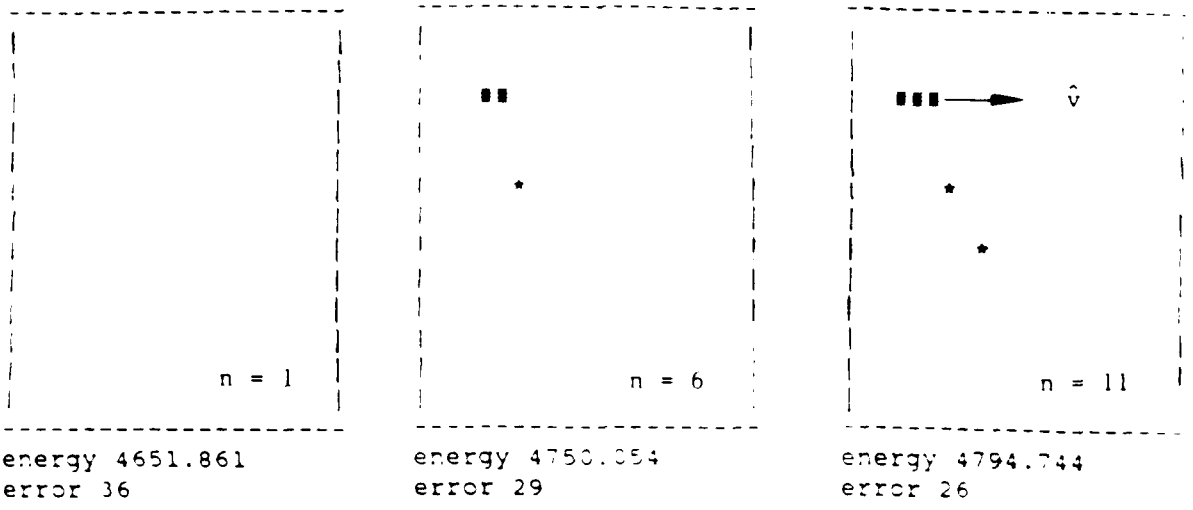


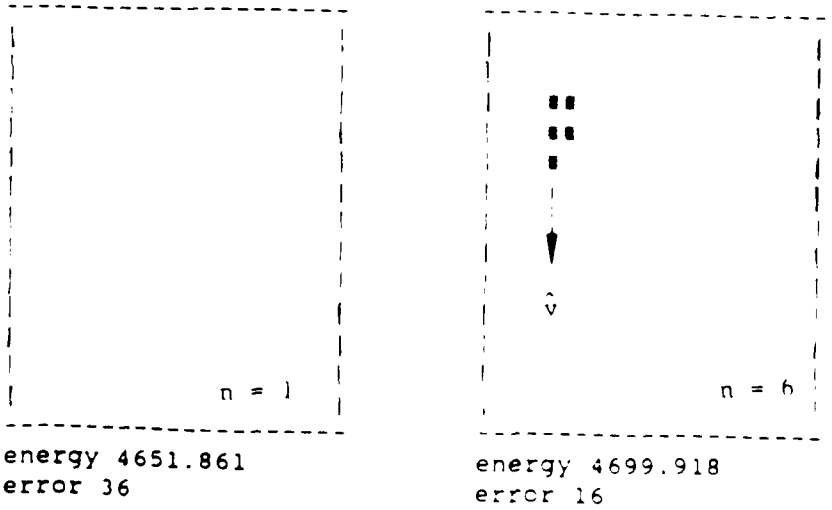
Figure 3.8 Successive Cluttered Images Containing Object over Cluttered Background, with Additive Noise.



(a) velocity match ($\hat{v}$= (1,1))



(b) velocity mismatch ($\hat{v}$ = (0,1))



(c) velocity mismatch ($\hat{v}$ = (1,0))

Figure 3.9 Object Recovery

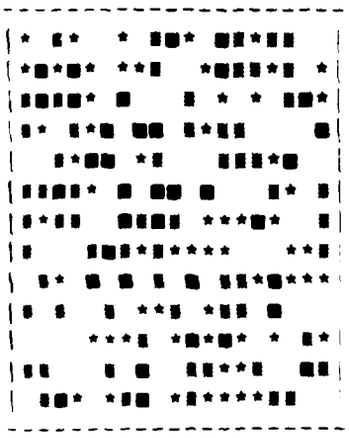
each noise source. The cluttered images are shown in Figure 3.10 and object recovery is shown in Figure 3.11. One object pixel is in error and there are three pixel artifacts.

3.4 Parallel Implementations

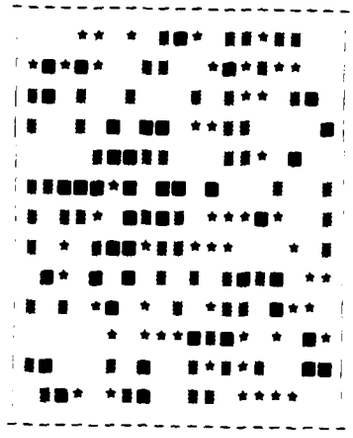
The randomized (Monte Carlo) procedure we have used for finding MAP solutions to the Bayesian estimation problem can be accelerated significantly by using a parallel architecture for implementation. If a processor is assigned to each pixel, for example, then the processing time will be reduced by a factor of n/k, where n is the total number of pixels and k is the chromatic number of the lattice. The chromatic number is equal to the minimum number of colors needed to color the sites of the lattice in such a way that no two neighbors have the same color. As an example, in our 4-connected neighborhood used for experimentation the chromatic number is 2.

To be more specific, consider solving for the MAP estimate by minimizing (3.8) using a massively parallel architecture such as the Connection Machine (Hillis, 1985), which is a Single Instruction Multiple Data (SIMD) array processor consisting of 256,000 processing units. Each unit has a single-bit arithmetic/logical unit and about 4k bits of storage and is organized in a 4-connected lattice that is 512 elements square. At each cycle of the machine, which we assume here to have a duration of one microsecond, an instruction is executed by each processor and a single bit is transmitted to its neighbors. This means that the first-order Markov field we have assumed here can be efficiently implemented.

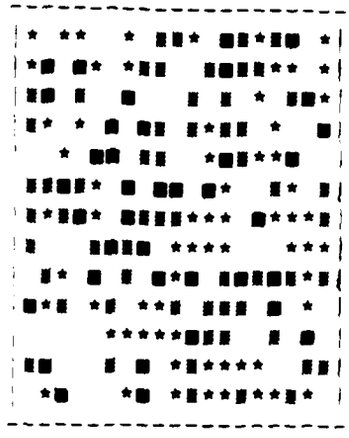
phi:k= 1



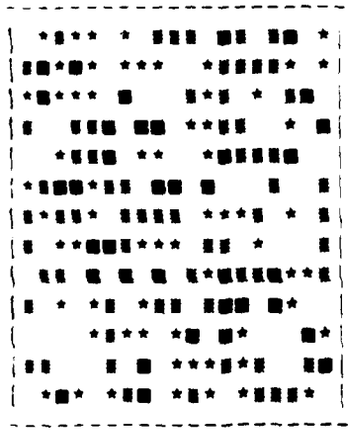
phi:k= 3



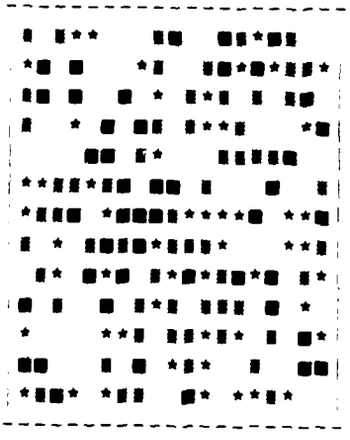
phi:k= 5



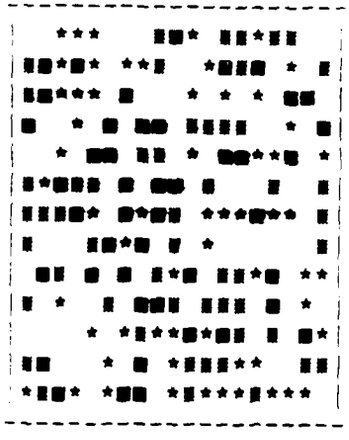
phi:k= 2



phi:k= 4



phi:k= 6



background(btru)

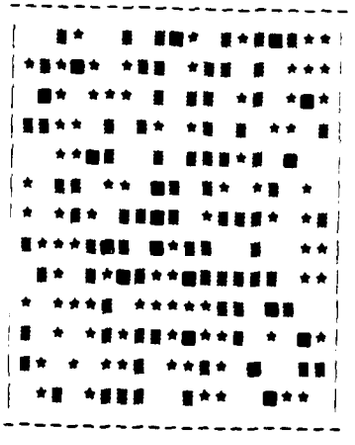
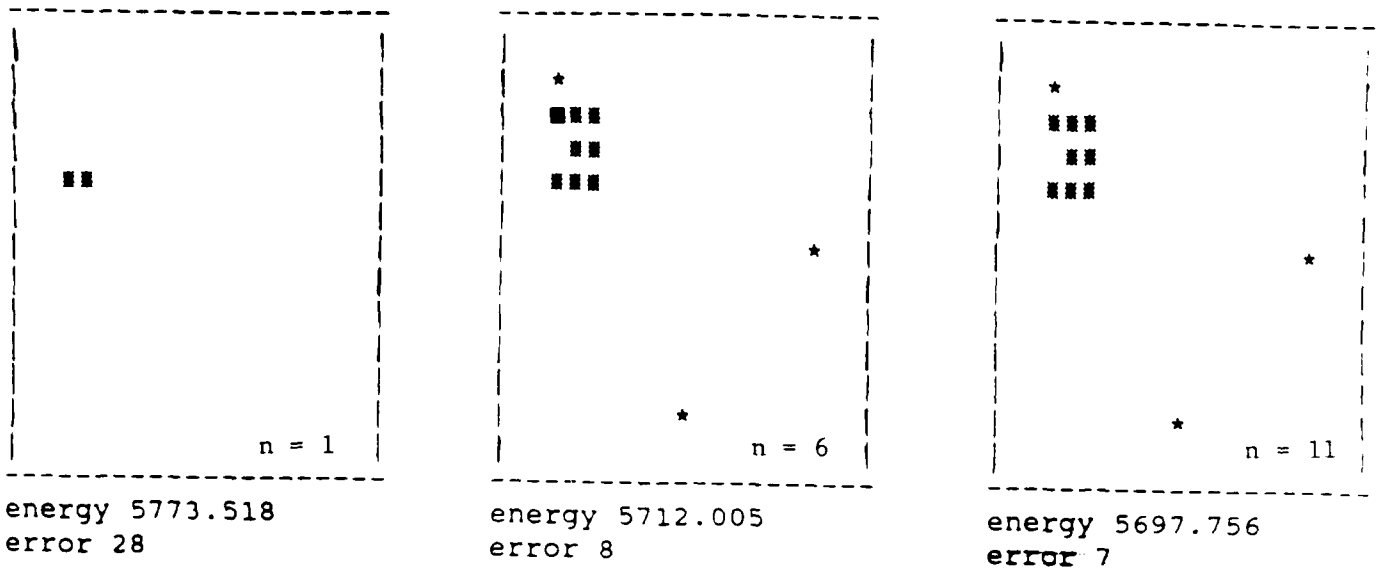


Figure 3.10 Cluttered Images and Background with Increased Noise (σ_B = 1, N_0 = 1, N_1 = 0.25).



energy 5773.518
error 28

energy 5712.005
error 8

energy 5697.756
error 7

Figure 3.11 Object Recovery with Increased Noise.

We assume that 16 cycles of a single-bit processor are required to perform a single 16-bit addition, subtraction or comparison; 256 cycles are required to perform multiplication or division; 512 cycles are required to generate a uniformly-distributed random variable; and 16 cycles are needed for memory transfer. We also assume a chromatic number of two, an overhead factor of two, 20 global iterations,and six sequential images to be analyzed. The resulting cycle count for a solution is 1,699,840 which yields a solution time of about 1.7 seconds. Note that for n pixels the number of processors required is 2 n v where v is the number of possible velocity combinations to be checked.

4. SUMMARY AND CONCLUSIONS

This research has been focused on the moving target detection and tracking problem using image analysis techniques. We were particularly interested in a severely cluttered environment in which high degrees of background and foreground noise were present. A probabilistic approach has been developed using an integrated Bayesian approach. Prior knowledge is employed using a Markov random field model in the form of a Gibbs distribution. Using probabilistic noise models, the aposteriori distribution is also Gibbs. No knowledge of the target is presumed, other than a local statistical model involving an arbitrarily-constructed potential function which then specifies the prior distribution.

The framework we use has been developed recently by several researchers and has yielded promising results for image segmentation, boundary detection and clutter rejection. In this research we have, so far as we know, extended the framework to include motion analysis for the first time. We develop both a batch processor and a recursive processor; the recursive processor yields superior results since the posterior conditional distribution is theoretically known exactly. The batch processor, on the other hand, is a smoother, and yields ordinary least-squares solutions.

We have developed parallelizable algorithms for computing the maximum aposteriori estimate of the image field containing the target. These algorithms are of the Monte Carlo type and are closely related to the Metropolis algorithm. Numerical results are presented to demonstrate the potential of this approach in solving very difficult target tracking

problems in an extremely noisy environment. We speculate that the results are close to the optimum in terms of accuracy in recovery of the unknown moving target.

Parallel implementation is discussed and an example given to demonstrate feasible solution times on current array processing hardware.

One of the most critical aspects of the problem is how to choose the system parameters to optimize performance. This is of particular importance when autonomous operation is required, as in SDI. Nonlinear optimization is required, once the interaction effects between algorithm parameters and measurable image parameters (such as statistical variation, correlation, etc) is worked out. This problem is virtually untouched in the literature and would be a suitable topic for Phase II research.

REFERENCES

Abend, K. "Compound Decision Procedures for Unknown Distributions and for Dependent States of Nature," in Pattern Recognition pp. 207-249, L. Kanal, ed. Thompson Book Co, Washington, DC (1968).

Adiv, Gilad, (1984), "Determining 3-D Motion and Structure from Optimal Flow Generated by Several Moving Objects," COINS Technical Report 84-07.

Bruss, A.R. and Horn, B.K.P., (1981) Passive Navigation, MIT A.I. Memo 662.

Cohen, F.S. and Cooper, D.B., (1984), "Simple Parallel Hierarchical and Relaxation Algorithms for Segmenting Noncausal Markovian Random Fields", Brown University Laboratory for Engineering Man/Machine Systems. Tech. Report LEMS-7.

Elliot H., Derin R., Christi R. and Geman D., (1983), "Application of the Gibbs Distribution to Image Segmentation," University of Massachusetts Technical Report.

Feller, W., (1950), "An Introduction to Probability Theory and its Applications, "Vol. I, Wiley, New York.

Geman S. and Geman D., (1984), "Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images," IEEE Trans. Pattern Analysis and Machine Intelligence 6, 721-741.

Grenander, U., (1984), "Tutorial in Pattern Theory," Division of Applied Math., Brown University.

Habibi, A., (1972), "Two Dimensional Bayesian Estimation of Images", Proc. IEEE 60, 878-883.

Hansen, A.R. and Elliot H., (1982), "Image Segmentation using Simple Markov Field Models," Comp. Graphics and Image Proc. 20, 101-132.

Hillis, D. (1985), "The Connection Machine", PhD Thesis, MIT Dept. of Electrical Engineering and Computer Science.

Hildreth, Ellen C., (1984), "Computations Underlying the Measurement of Visual Motion", Massachusetts Insitute of Technology, A.I. Memo No. 761.

Kemeny, J.G. and Snell, J.L., (1960), "Finite Markov Chains", Van Nostrand, New York.

Legters, G.R., and Young, T.Y. (1982), "A Mathematical Model for Computer Image Tracking", IEEE Trans. Pattern Anal. and Mach. Intell. Vol. PAMI-4, No. 6, pp. 583-594.

Marroquin, Jose Luis, (1985), Probablistic Solution of Inverse Problems," Massachusetts Institute of Technology.

Marroquin, J., S. Mitter and T. Poggio, (1986), "Probabilistic Solution of Ill-Posed Problems in Computational Vision," MIT Artificial Intelligence, Intelligence Lab.

Metropolis, N., Rosenbluth, A.W., Teller, A.H. and Teller, E. (1953), "Equations of State Calculations by Fast Computing Machines", J. Chem. Phys., Vol. 21, pp. 1087-1091.

Miller, K.S., Raghavan, R. and Rochwarger, M.M. (1985), "Invariance in Moving Target Detection", IEEE Trans. Inf. Theory, Vol. IT-31, No. 1, pp. 69-79.

Morozov, V.A., (1984) Methods for Solving Incorrectly Posed Problems, Springer-Verlag, New York.

Morroquin, J., Mitter, S., and Poggio, T. (1987), "Probabilistic Solution of Ill-Posed Problems in Computational Vision", J. Amer. Stat. Assoc., Vol. 82, No. 397, pp. 76-89.

Nahi, N.E., and Assefi, T., (1972), "Bayesian Recursive Image Estimation," IEEE Trans. on Computers 21, 734-738.

Poggio, T. and Koch C., (1984), "Analog Networks: A New Approach to Neutral Computation," Artificial Intelligence Lab. Memo 783, MIT.

Poggio, T. and Torre, V., (1984), "Ill-Posed Problems and Regularization Analysis in Early Vision," A.I. Memo 773, MIT.

Poggio, T., Torre, V. and Koch C., (1985), "Computational Vision and Regularization Theory," Nature.

Reed, L.S., Gagliardi, R.M. and Shao, H.M. (1983), "Application of Three-Dimensional Filtering to Moving Target Detection", IEEE Trans. Aero and Electronic Systems, Vol. AES-19, No. 6, pp. 898-904.

Terzoupulos, D., (1984), "Multiresolution Computation of Visible-Surface Representations," Ph.D. Thesis, Dept. of E.E. and C.S., MIT.

Terzoupulos, D., (1985), "Integrating Visual Information for Multiple Sources for the Cooperation Computation of Surface Shape" to appear in From Pixels to Predicates: Recent Advances in Computational and Robotic Vision, ed. A. Pentland, Ablex.

Ullman, S., (1981), "Analysis of Visual Motion by Biological and Computer Systems", Computer 14.

Williams, T.D., (1981) "Computer Interpretation of a Dynamic Image from a Moving Vehicle," Ph.D. Dissertation (TR 81-22), Computer and Information Science Dept., Univ. of Mass.

END

DATE

FILMED

DTIC

JULY 88